5-2011

Gas-Phase Thermochemistry of Non-Protein Amino Acids by the Extended Kinetic Method

Vincent James Yannello
College of William and Mary

Gas-Phase Thermochemistry of Non-Protein Amino Acids by the Extended Kinetic Method

A thesis submitted in partial fulfillment of the requirement
for the degree of Bachelor of Science in Chemistry from
The College of William and Mary

by

Vincent James Yannello

Accepted for _____

_____
Dr. John Poutsma, Director

_____
Dr. Christopher Abelt

_____
Dr. Rex Kincaid

Williamsburg, VA
May 3, 2011

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

First of all, I would like to thank Dr. Poutsma for all of the help and opportunities he has given me the past few years, and all of the help, support, and patience in the making of this thesis. I would also like to thank my parents, sisters, and other family members for their support in all of my work. I would like to thank my friends, whose support and friendship have helped me through everything. Lastly, I would like to thank all of the members of Ionlab, for their encouragement, support, and laughter through all of our research experiences.

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

In this study, we have been examining the gas-phase acidities of two amino acid homologues, homocysteine and homoserine, and the proton affinity of another amino acid homologue, homohomoserine. Homocysteine, homoserine, and homohomoserine are non-protein amino acids (NPAA) that have very similar structures to the protein amino acids (PAA) cysteine and serine. The gas-phase acidities of both cysteine and serine are well known and an analysis of the gas-phase basicity of their homologues has yielded information regarding the effects of small changes in structure on a basic physical property. Similarly, the proton affinities of serine and homoserine are well-known, so investigating the proton affinity of the next homologue has yielded more information as well. To do this experiment, we used a triple quadrupole mass spectrometer with an electrospray ionization source and utilized the kinetic method. In this method, the gas-phase acidity of the compound of interest is determined by comparing it to other compounds of known gas-phase acidity, called reference acids, inside the mass spectrometer. A proton-bound dimer is formed and is collided with an inert gas at various energies to obtain a kinetic plot. This is repeated for multiple reference acids and an intersection is found. This intersection is the gas-phase acidity of the compound of interest. A similar analysis was done for the proton affinity of homohomoserine. The final values obtained were 1394.6 ± 10 kJ/mol for the gas-phase acidity of homoserine, 1392.8 ± 10 kJ/mol for the gas-phase acidity of homocysteine, and 968.7 ± 10 kJ/mol for the proton affinity of homohomoserine.

Introduction:

Mass Spectrometry:

Mass spectrometry is an analytical technique that can be used to separate ions from complex mixtures by their mass to charge (m/z) ratio. The first modern mass spectrometer was developed by Arthur Jeffrey Dempster, in 1918. His mass spectrometer had accuracy over 100 times greater than all previous mass spectrometers, and his design and theory are the bases for mass spectrometers used today. Figure 1 shows a diagram of his mass spectrometer. The basic principle behind his spectrometer was to introduce the source ions in the chamber marked by the letter "G." Here they were accelerated into the analysis chamber labeled "A." The curvature of the path the ion travels is determined by the ion's mass, charge, and the potential difference applied across the system. Therefore, given a set potential difference, a single m/z ratio can be isolated and analyzed, in this case, in the chamber labeled "E."[1]

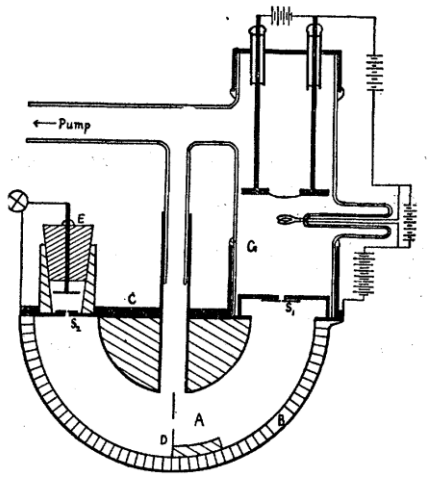


Fig. 1.

Figure 1 – Dempster's First Mass Spectrometer[1]

Modern mass spectrometers use the same basic theory of separation by m/z ratio, though the methods used today are far more complex and varied. As mass spectrometers analyze ions in

the gas-phase by the m/z ratio, the system under which all testing is performed must be under a very high vacuum. Two vacuum pumps are used, the first of which is called a rotary vane mechanical pump. This pump lowers the pressure inside the machine to levels in which the second pump, called the turbo pump can then operate. This pump will decrease the pressure even more, typically in the neighborhood of 1×10^{-5} torr. This low pressure allows the large number of ions produced in the mass spectrometer to make their way through the analyzer to the detector without colliding with another molecule or ion and losing its charge. The detector returns a signal which is interpreted by the computer, and a spectrum is printed on the screen. The spectra are typically shown with the m/z ratio on the x-axis, while the relative intensity is shown on the y-axis, scaled from zero to 100%. An example spectrum is shown in Figure 2.

There are two integral portions of the mass spectrometer which differ depending on the experiment being performed: the source and the analyzer. Sources exist for gas, liquid, and solid phase samples. For gaseous samples, the most common techniques are electron ionization and chemical ionization. In electron ionization, the gas is sent through a stream of electrons produced by a powerful voltage. This technique is called a "hard" ionization source, since the ions produced are likely to fragment due to the large amount of energy being imparted to them. In chemical ionization, the ions are created in the gas phase by a collision with another ion already present in the source, typically obtained from ionizing methane or ammonia using electron ionization. This form of ionization is considered "soft" due to the fact that the ions produced do not tend to fragment before being analyzed.

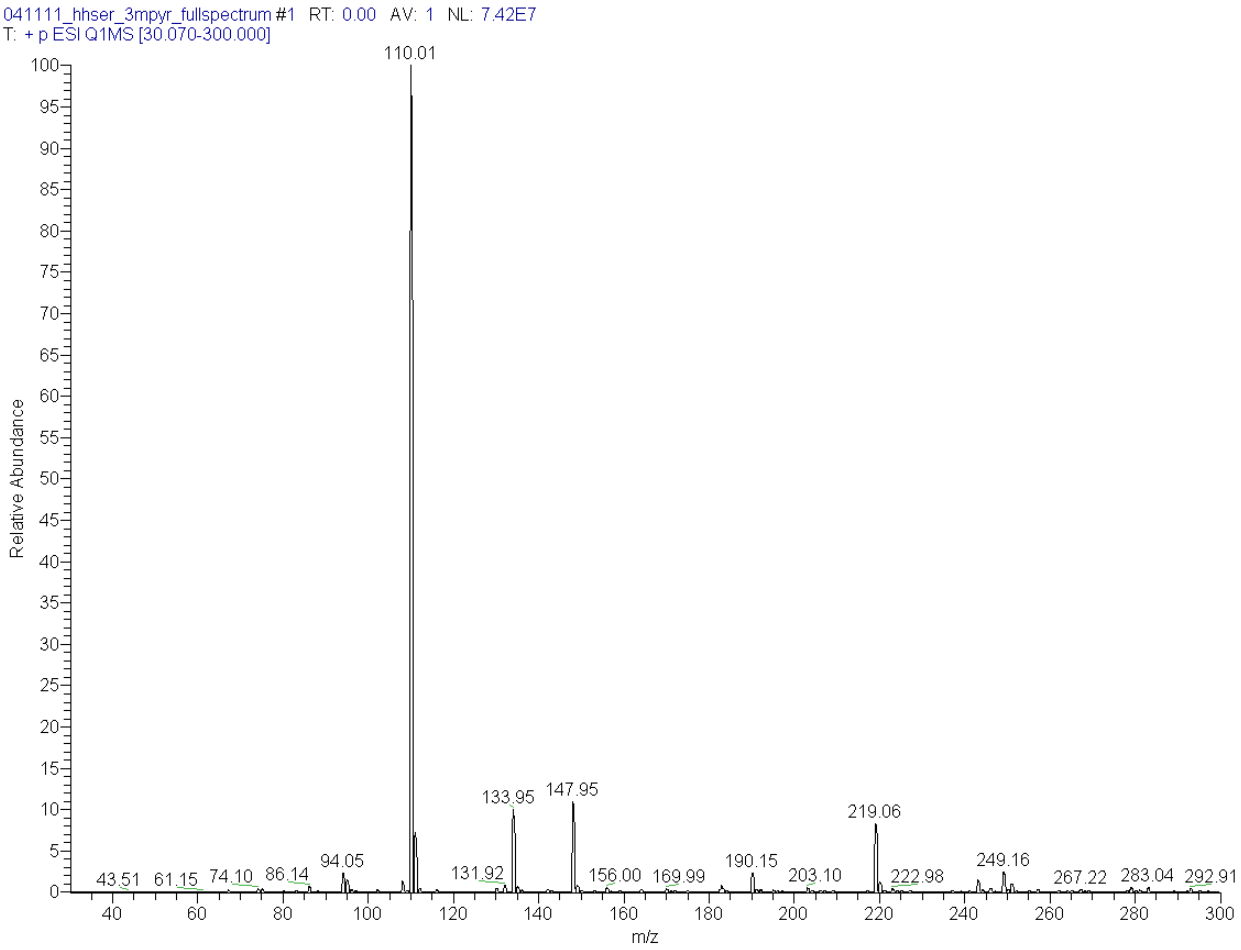


Figure 2 – Example Mass Spectrum

For a liquid sample, the most common ionization source is called an electrospray ionization (ESI) source. In an ESI source, the compound of interest is dispersed into a fine aerosol from an electrically-charged needle, which creates small charged droplets. Ions are desolvated by a flow of inert gas, such as nitrogen, and by passing them through a heated capillary. Eventually, the charged molecule will be left in the gas phase after which it can be analyzed.[2] Like chemical ionization, this is a "soft" ionization technique, and one of the best "soft" ionization sources for a liquid sample.

For a solid sample, a matrix-assisted laser desorption ionization (MALDI) source is typically used, in which a laser is shot at the solid sample, which promotes vaporization and

ionization of the sample. This technique is also a "soft" ionization technique, and is typically used to analyze many biomolecules, especially rather large ones, such as proteins, polymers, and sugars.[3] Along with all of the previously mentioned sources, many more exist and are currently under development, such as atmospheric pressure chemical ionization, inductively coupled plasma, among others.[4]

The analyzers used in mass spectrometer differ greatly as well. The first modern mass spectrometer made by Dempster is called a sector instrument. All sector instruments physically bend the path which the ions travel through the analyzer and are based on either electric or magnetic fields. The larger the m/z ratio is, the less the ion will be deflected, and similarly, the smaller the m/z ratio is, the more the ion will be deflected. Using this information, it is possible to determine the exact m/z ratio given set potential differences in the spectrometer. A second kind of analyzer, typically paired with a MALDI source, is called a time of flight (TOF) analyzer.[3] In a TOF analyzer, the ions are accelerated through a constant and known potential, and the time it takes to reach the detector is used to determine the m/z ratio. As such, a smaller m/z ratio will be detected sooner than a larger m/z ratio.

Two additional kinds of analyzers are commonly used today, and they are called the quadrupole mass analyzer and the ion trap mass analyzer. Both work on the same basic principle of applying varying radio-frequency (RF) voltages to a set of electrodes, which gives ions trajectories based on their m/z ratio. In a quadrupole mass filter, the voltages can be setup such that only a specific m/z ratio will successfully be able to traverse the analyzer without colliding into one of the rods due to a destabilized trajectory. All of the ions with the correct m/z ratio are passed through the system to the detector. Thus, a quadrupole mass filter can separate a mixture of ions by their m/z ratio. One very common use of the quadrupole mass filter is called a triple

quadrupole setup. In this system, three quadrupole mass filters are put in a line, so that multiple isolations can occur. Typically, the first and last quadrupoles will select and/or scan the mass ranges, while the second quadrupole is a collision chamber. By doing this, an MS-MS setup is created that can analyze compounds in ways that could not be analyzed before.

The ion trap mass analyzer works by the same principle as the quadrupole mass analyzer, by having RF voltages applied to a set of electrodes However, unlike the quadrupole filter, the ion trap physically traps the ions. To generate a mass spectrum, ions of a certain m/z are ejected by causing their trajectory to become unstable in the z-direction. As with the ionization sources, there are a large number of other analyzers, such as a fourier transform ion cyclotron resonance analyzer.[5] An example setup of an electrospray ionization source and an ion trap mass analyzer is shown in Figure 3 below, including typical values used in an experiment using this setup.

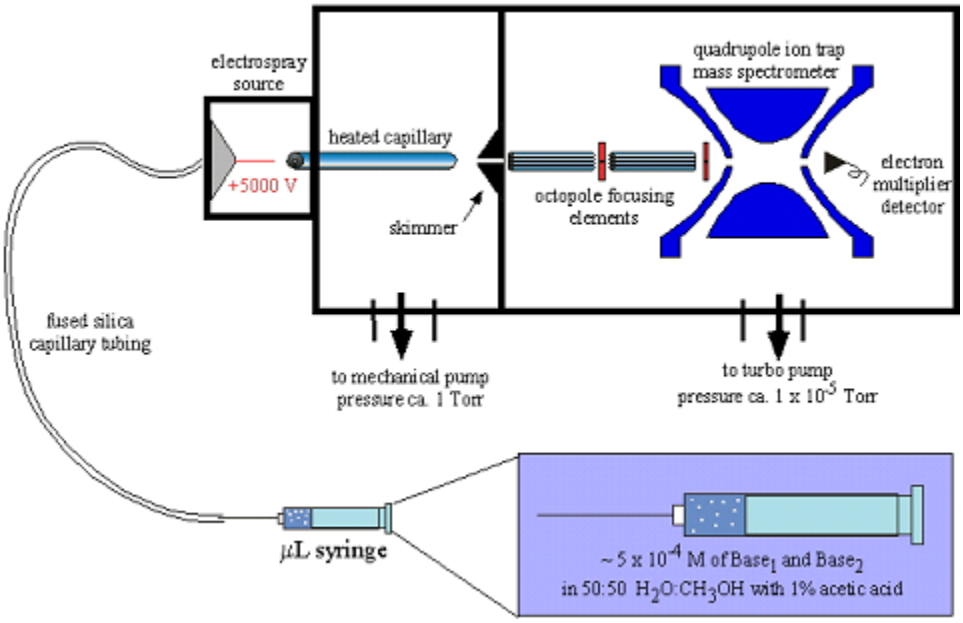


Figure 3 – Example Mass Spectrometer Setup

As mentioned with triple quadrupole mass analyzers, tandem mass spectrometry has become a very useful analytical tool. By placing multiple quadrupole analyzers in a row, many

new experiments can be run to further analyze the compound.[3] Some examples are parent ion and product ion detection modes. In the parent ion mode, the first quadrupole scans a m/z ratio range set by the experimenter, the second quadrupole acts a collision cell, and the third quadrupole detects one specific m/z ratio. In the end, the m/z parent ions that produce a certain product ion when collided are detected. Product ion mode works in a similar manner, except the first quadrupole is set to one m/z ratio, while the third quadrupole scans all m/z ratios. This yields what is called a fragmentation spectrum which can show how a certain ion fragments upon being collided with a given energy. An example of a fragmentation spectrum is shown in Figure 4. Product ion scans can also be performed using an ion trap analyzer, though the method by which is performs them is slightly different. The one advantage to an ion trap analyzer is that one ion trap analyzer can perform the job of "n" quadrupole analyzers, where n is any number greater than or equal to one.[2]

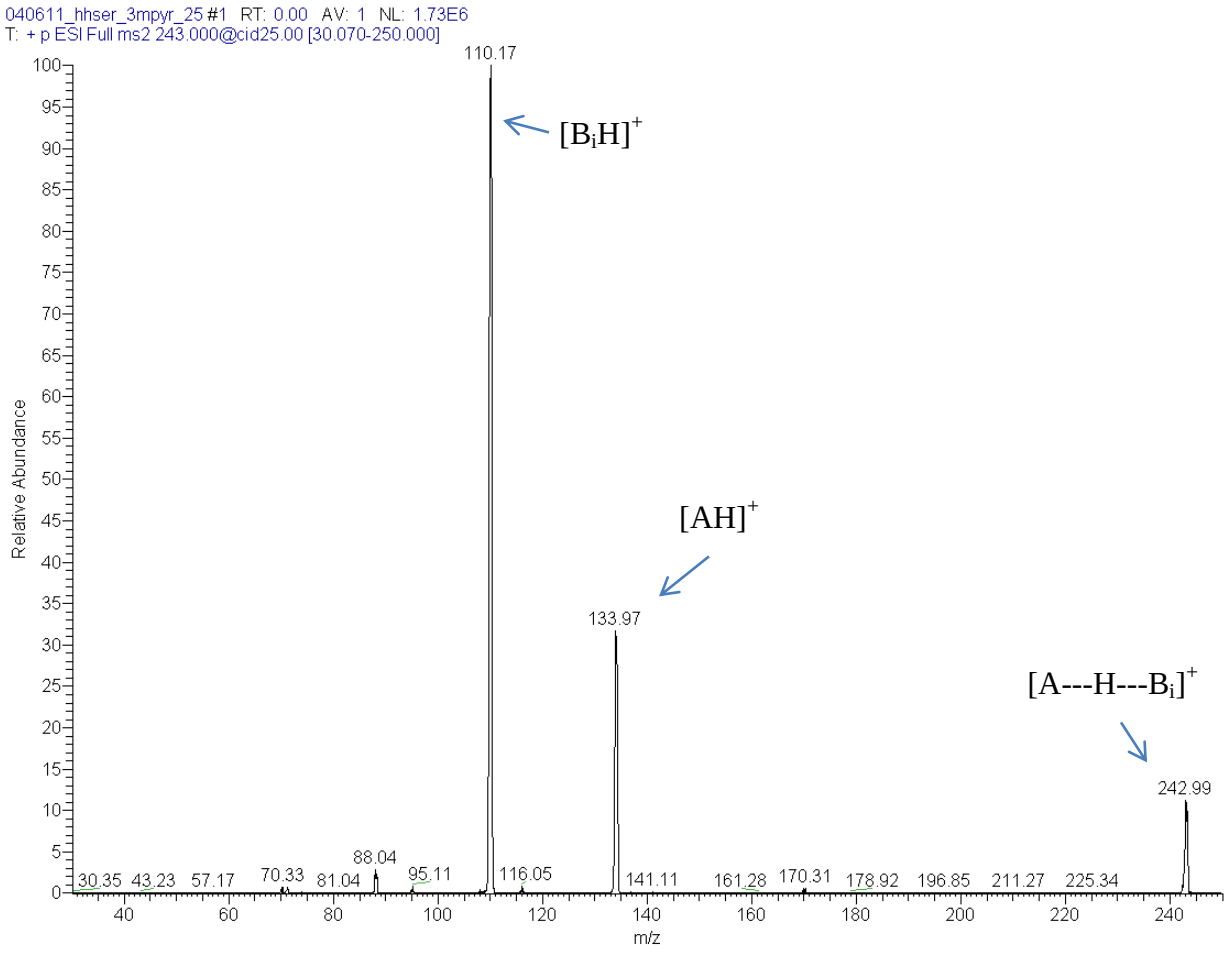


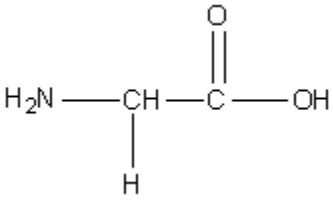
Figure 4 – Example Fragmentation Spectrum

The applications of a mass spectrometer are numerous and have been developed extensively ever since the introduction of these techniques. One of the first applications of the mass spectrometer was the determination of isotopes in a sample. As the isotopes will typically differ by one or two atomic mass units, most other techniques have a hard time distinguishing between them. The mass spectrometer, however, can resolve these differences with ease, and is used to find the ratio of isotopes in samples. As mass spectrometers are very sensitive, they can typically determine small amounts of rare isotopes in samples.[6] Due to their sensitivity, mass spectrometers can also be used in determining trace molecules in a large variety of samples, and

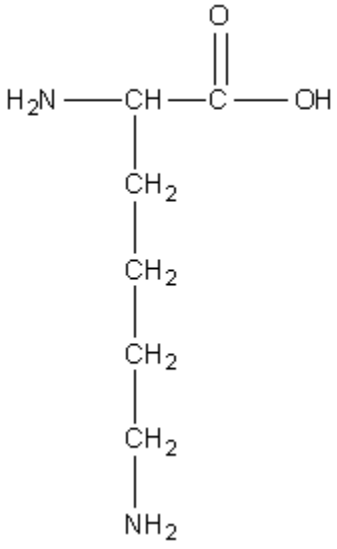
especially gases.[6] Mass spectrometry also has many uses in studying proteins and their structures.[7]

Proteins and Amino Acids:

Proteins are made up of a linear chain of molecules called amino acids, which are small molecules consisting of an amine group on one end (called the N-terminus) and a carboxylic acid on the other end (called the C-terminus). The amino acids that make up the large portion of proteins seen in humans are α-amino acids, because there is only one (-CH) group separating the amine and the carboxylic acid functional groups. Attached to this intermediary (-CH) group is what is called a side chain (R). For each amino acid, this side chain varies from the simplest amino acid glycine (R =H) to the more complicated lysine (R=CH_2-CH_2-CH_2-CH_2-NH_3) shown in Figure 5. Most proteins are made up of only twenty of the α-amino acids, called protein amino acids (PAAs). Other naturally occurring amino acids are classified as non-protein amino acids (NPAAs).



Glycine



Lysine

Figure 5 – Structures of Glycine and Lysine

Proteomics is the study of the structure of composition of proteins. As most proteins are very large compounds (tens of thousands of amu) it can be difficult to determine chemical composition and thermodynamic properties. However, as proteins are easily charged, a mass spectrometer can be used to analyze them. As the proteins will easily fragment upon collision, their fragments can also be analyzed in a tandem mass spectrometer to give more clues to the structure of the protein. Though starting from a protein and slowly breaking it apart is one method to get more information about the protein, studying the amino acids which make up the protein is an equally interesting and rewarding process.

In proteins, amino acids interact with each other through their side chains. Of all of the protein amino acids, some side chains are hydrophobic, some are hydrophilic, some are acidic, and some are basic. Many of the protein amino acids can also be involved in hydrogen bonding, both as donors and receptors. Through these various interactions, the protein "folds" to form the lowest energy structure between these interactions. When a protein is made, it is possible that a wrong amino acid may be incorporated into the protein. This can result in the protein being inactivated due to its inability to perform its primary function anymore. The easiest form of misincorporation is taking a non-protein amino acid homologue instead of the protein amino acid itself. This is due to the similar structure between the two amino acids. Though the structures may be similar, the functioning of the two may be very different.

In the research lab, many analyses are done on amino acids and their homologues. The kinetic method is performed on compounds to determine their proton affinity and gas-phase acidity. Additionally, hydrogen-deuterium (H/D) exchange experiments are done to determine experimentally which hydrogen atoms are the most acidic or basic. Finally, theoretical

calculations are typically done on the molecules as well to get further confirmation for the values obtained experimentally.

In studying amino acids in the gas phase, many thermodynamic properties can be determined. Two of the most important are proton affinity (PA) and gas-phase acidity (GA). The proton affinity is defined to be the negative of the change in enthalpy of the protonation reaction in the gas phase, shown below:[8]

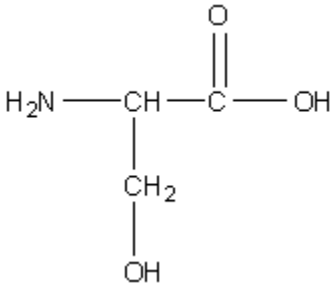
$$A + H^+ \rightarrow AH^+ \qquad\qquad PA = -\Delta H$$

In a similar manner, the gas-phase acidity of a compound is defined to be change in enthalpy of the deprotonation reaction of the neutral compound, shown below:

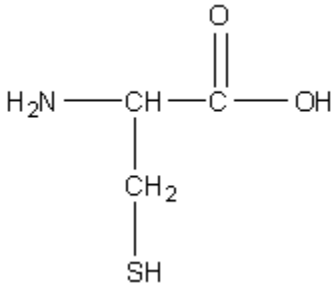
$$AH \rightarrow A^- + H^+ \qquad\qquad GA = \Delta H$$

Both of these quantities can be used to determine information about not only the structure of the molecule in the gas phase, but also give clues to the structure in solvated phases and relative preferences for protonation and deprotonation. Though less common, the sodium affinity and the potassium affinity of amino acids can also be found by using a mass spectrometer. Both of these values have implications to the binding factor of various positive ions to amino acids while in the protein.[9]

Two very interesting protein amino acids are serine and cysteine, shown in Figure 6. These two amino acids are extremely similar, differing only by one atom on their side chain; an oxygen in serine, and a sulfur in cysteine.



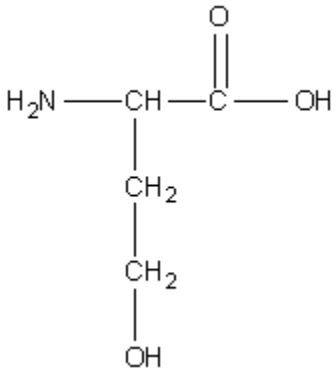
Serine



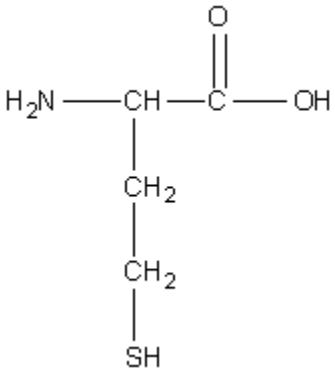
Cysteine

Figure 6 – Serine and Cysteine

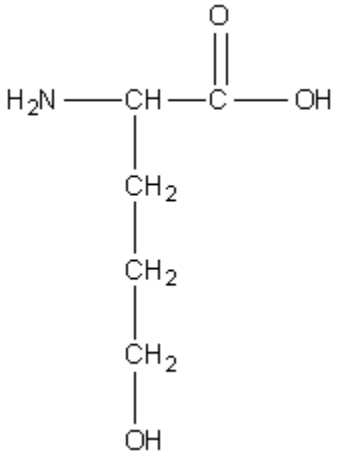
Both serine and cysteine have non-protein amino acids homologues in which the side chain differs by one additional (-CH_2) group. These two amino acids are called homoserine and homocysteine, and are shown along with homohomoserine and homohomocysteine in figure 7. Homohomoserine and homohomocysteine differ from homoserine and homocysteine also by an additional (-CH_2) group. When the protein is being made, a misincorporation of any of these four non-protein amino acids could have a large effect on the functioning of the protein, or it could have little to no effect.



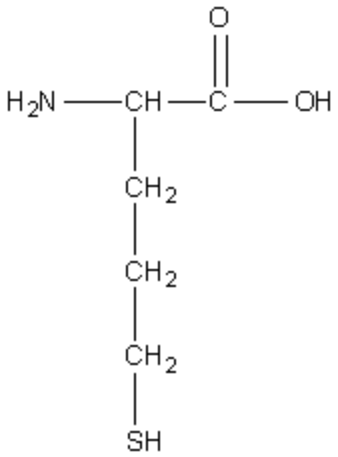
Homoserine



Homocysteine



Homohomoserine



Homohomocysteine

Figure 7 – Structures of serine and cysteine homologues

The easiest way to determine this effect would be to look at the intrinsic thermodynamic properties of all six compounds. Though the proton affinity and gas-phase acidity of serine and cysteine are well-known in the literature[15], little research has been done to determine these values for the four homologues. In investigating how the side chain length affects these values, it is possible to determine whether or not this side chain length would be important to the function of cys and ser-contaning proteins.

Another interesting property of cysteine and its homologues is the acidity of the side chain group. All of the other protein amino acids have gas-phase acidity values determined both theoretically and experimentally that confirm that the deprotonation in the gas phase occurs on the C-terminus end of the amino acid. However, since cysteine's (-SH) group is also very acidic, recent calculations and experimental evidence have had conflicting results as to the site of deprotonation.[10,11] Not surprisingly, the same evidentiary clash has occurred for cysteine's two homologues, homocysteine and homohomocysteine. However, unlike cysteine, homocysteine

and homohomocysteine do not have gas-phase acidity values computed experimentally yet. These two reasons have motivated this work in investigating the proton affinity and gas-phase acidity for homoserine, homocysteine, homohomoserine, and homohomocysteine.

The Kinetic Method:

The Cooks kinetic method is used to determine these thermodynamic quantities. The kinetic method works by the formation and resulting dissociation of a proton-bound heterodimer between two compounds [A --- H^+ --- B_i]· A is the compound with unknown thermochemical property, and B_i is the i^{th} member of a set of reference bases with known thermochemical properties. When this heterodimer dissociates, the H^+ will stay attached to either A or B_i, and the relative amounts of each protonated product will be in a certain ratio which is approximately equal to the ratio of the respective rates of dissociation. This ratio is measured by tandem mass spectrometry. Using transition state theory, the natural logarithm of the ratio of the rates of dissociation is given by the following formula

$$\ln\left(\frac{k_A}{k_i}\right) = \ln\left(\frac{Q_A^*}{Q_i^*}\right) - \frac{\varepsilon_A^\circ - \varepsilon_i^\circ}{RT_{eff}}$$

with k_A and k_i defined to be the rates of dissociation of the proton-bound dimer to the respective ions, shown below.

$$[A - H - B_i]^+ \rightarrow AH^+ + B_i \qquad rate = k_A$$

$$[A - H - B_i]^+ \rightarrow A + B_iH^+ \qquad rate = k_b$$

Q_A* and Q_i* are the corresponding partition functions of the activated complexes, and ε_A° and ε_i° are the corresponding activation energies. R is the gas constant, and T_{eff} is the effective temperature of the system.[12] In the kinetic method, the proton-bound dimer is "activated" by imparting a certain amount of energy into this dimer and then colliding it with some inert substance, normally an inert gas. The difference in activation energies is assumed to be the difference in proton affinities of A and B_i, written as

$$PA(A) - PA(B_i) = \Delta PA$$

which is based off of the presumption that the dissociations occur without reverse activation energy.

The partition function can also be rewritten as a difference between an apparent protonation entropy difference between A and B_i, or

$$\ln\left(\frac{Q_A^*}{Q_i^*}\right) = -\frac{\Delta S(A) - \Delta S(B)}{R}$$

Combining both of these approximations yields the following equation

$$ln\left(\frac{k_A}{k_i}\right) = \frac{\Delta PA}{RT_{eff}} - \frac{\Delta\Delta S}{R}$$

with $\Delta\Delta S = \Delta S(A) - \Delta S(B_i)$.[12] The standard kinetic method assumes that the entropy term is approximately equal to zero. This is generally a good assumption if the molecules being tested

are of about the same size and have a similar structure. In the cases of more complicated molecules, the entropy term cannot be ignored, and the extended kinetic method must be used to get an accurate value of the proton affinity.[13] The key to the extended method is in the usage of different activation energies. Since each reference base has a different proton affinity, the ratio of its dissociation rate to the compound of interest will be different for each reference base. If for each reference base, the ratio of peaks is scanned over a large range of effective temperatures, it is possible to get a plot showing the change in the ratio of the dissociation rates compared to the difference between the reference bases proton affinity to the average proton affinity of all used reference bases. Lines are drawn for each effective temperature, and the intersection between all of them is the isothermal point. This point shows the estimated value of the proton affinity for the compound.[13]

Though all of the previous analysis was done to show how to compute the proton affinity of a compound, a nearly identical analysis can be performed to determine the gas-phase acidity of a compound. The key difference is the dimer being investigated. To determine the gas-phase acidity of the compound, negative ions must be formed, and as such, a negative proton bound dimer is formed, of the form $[A \text{ --- } H \text{ --- } B_i]^-$. The analysis is exactly the same from here on out, save a sign change in front of the $\Delta GA/RT_{eff}$ term.

Experimental Procedure:

All experiments were run in a TSQ-QUANTUM ThermoFinnigan triple quadrupole mass spectrometer equipped with an electrospray ionization source. The solutions for the reference base/acid and the amino acid were initially prepared at 5×10^{-3} M, and then they were then subjected to a 10:1 dilution. For positive ion testing, the solutions were prepared in 49.5:49.5 water:methanol solvent with an additional 1% acetic acid added to aid in positive ion formation. For negative ion testing, the solutions were prepared in a 79.2:19.8 methanol:water solvent with an additional 1% ammonium hydroxide added to aid in the formation of negative ions. The solutions used for testing were made from mixing 1 mL of one reference base/acid with 1 mL of the amino acid. To determine the proton affinity, the mass spectrometer was operated in positive ion mode, while the mass spectrometer was run in negative ion mode to determine the gas-phase acidity.

Solutions were injected into the electrospray ionization source using a 500 μL Hamilton Gastight® syringe. The heated capillary was maintained at 160° C, while all other parameters corresponding to the electrospray source were set by the software to maximize the ion count of the dimer. A 4 kV charge is applied to the electrospray needle to produce charged droplets. As the droplets are propelled by the nitrogen sheath gas (8 arbitrary units), the solvent slowly evaporates leaving only the ions. Collision Induced Dissociation (CID) was used to fragment the dimer, with the argon collision gas pressure set to 0.3 arbitrary units.

Initially, the MS spectrum was evaluated to make sure that the dimer was present. Other ions that were expected were the base/acid ions, the amino acid ions, and the homodimers between the base/base or acid/acid, and the amino acid/amino acid. The heterodimer was isolated first, with no collision energy. Scans were taken as the total ion count observed for a length of

one minute to ensure adequately high levels of ions were counted. The collision energy was ramped in increments of five V, with scans performed at each increment. This continued until collision energy of 75 V was reached. As the collision energy goes up, the ion count for the dimer decreases, as more of the dimer fragments into the base/acid and amino acid molecules. The molecule with a higher signal has the higher apparent gas-phase basicity/gas-phase acidity. Eventually, the base peak for the collision spectrum is either the base/acid ion or the amino acid ion. Other ions which may have significant ion counts are secondary fragmentation of the acid/base or amino acid and are included in the total count when applying the kinetic method. In some spectra, more ions are detected in significant levels that are merely contamination of the solution. These generally do not affect the results of the kinetic method. After running the tests for all collision energies, the mass spectrometer was flushed with approximately 250 μL of either 49.5:49.5:1 methanol:water:acetic acid or 79.2:19.8:1 methanol:water:ammonium hydroxide for positive ion mode and negative ion mode, respectively.

Data was collected for each reference base a total of three times. The values used in the kinetic method for each collision energy were the averages of the three different days on which data were taken.

Results:

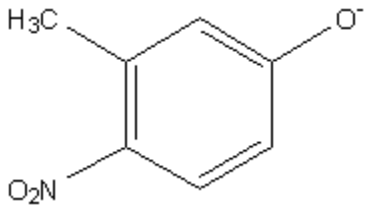
Thermochemical Data for Homoserine:

In investigating homoserine, the gas-phase acidity was the key value being observed. To determine the gas-phase acidity, a total of six reference acids were used. They are listed below in Table 1, with their m/z ratio of the negative ion, and their gas-phase acidity listed as well. In Figure 8, all six structures are given.

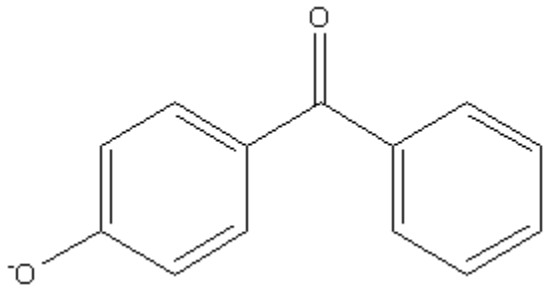
Table 1 – Reference Acids for Homoserine

Reference Acid	*Molecular Mass(g/mol)*	*Gas-Phase Acidity (kJ/mol)*
3-Methyl-4-Nitrophenol	153	1380
4-Hydroxybenzophenone	198	1393
3-Nitrophenol	139	1400
Chloroacetic Acid	94	1408
2-Fluorobenzoic Acid	140	1415
2,5-Dimethylbenzoic Acid	150	1420

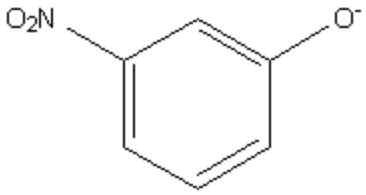
The m/z ratio for homoserine is 118, the dimer mass was easily calculated for each homoserine/reference acid pair. As 3-nitrophenol and 2-fluorobenzoic acid have m/z ratios very close to one another, care was taken to flush between all runs and to make sure that these two acids were not run in the same time frame. This was done to prevent any possible leftover masses being mistaken in the fragmentation spectra.



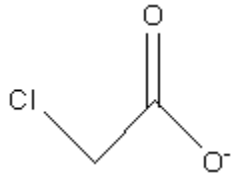
3-Methyl-4-Nitrophenol
m/z = 152



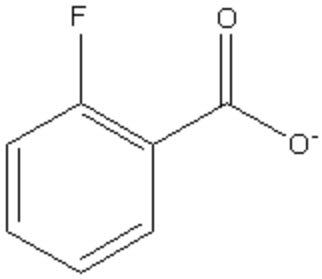
4-Hydroxybenzophenone
m/z = 197



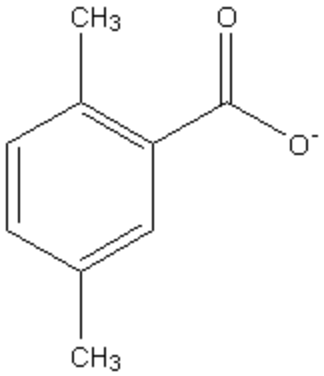
3-Nitrophenol
m/z = 138



Chloroacetic Acid
m/z = 93



2-Fluorobenzoic Acid
m/z = 139



2,5-Dimethylbenzoic Acid
m/z = 149

Figure 8 – Structures and m/z for deprotonated homoserine reference acids

In determining which reference acids would be used for this experiment, the theoretical acidity of the compound was used as an estimate of the true value. Reference acids with gas-phase acidities near this value were tested to see if they formed favorable dimers with homoserine. Of all of the reference acids tested, these six yielded the best dimer formation and fragmentation results.

To determine the gas-phase acidity of homoserine, two separate kinetic plots were drawn. A kinetic plot is one of two plots used in the kinetic method to accurately determine the gas-phase acidity of the compound of interest. Kinetic plot 1, shown in Figure 9, is a graph of

ln(relative intensity) vs. GA – $GA_{average}$ for each reference acid and at different activation energies. As the extended kinetic method requires multiple activation energies, energies between 0 and 75 arbitrary units were used. As a plot of all 16 collision energies would be difficult to interpret, and as some energies are not conducive to accurate fragmentation, only a few of the energies are plotted. A line of best fit is created for each set of collision energies across all reference acids. These lines should intersect at a single point, but typically do not do to experimental error.

To determine which energies will be plotted, a graph of the collision energies in arbitrary units vs. the effective temperature is drawn. From this, the area where the steepest incline is achieved is the best for determining the gas-phase acidity. The graph to determine the range of energies used for Homoserine is shown in Figure 10 below. For this experiment, a total of 5 energies were chosen within the range of 10 to 45 arbitrary units.

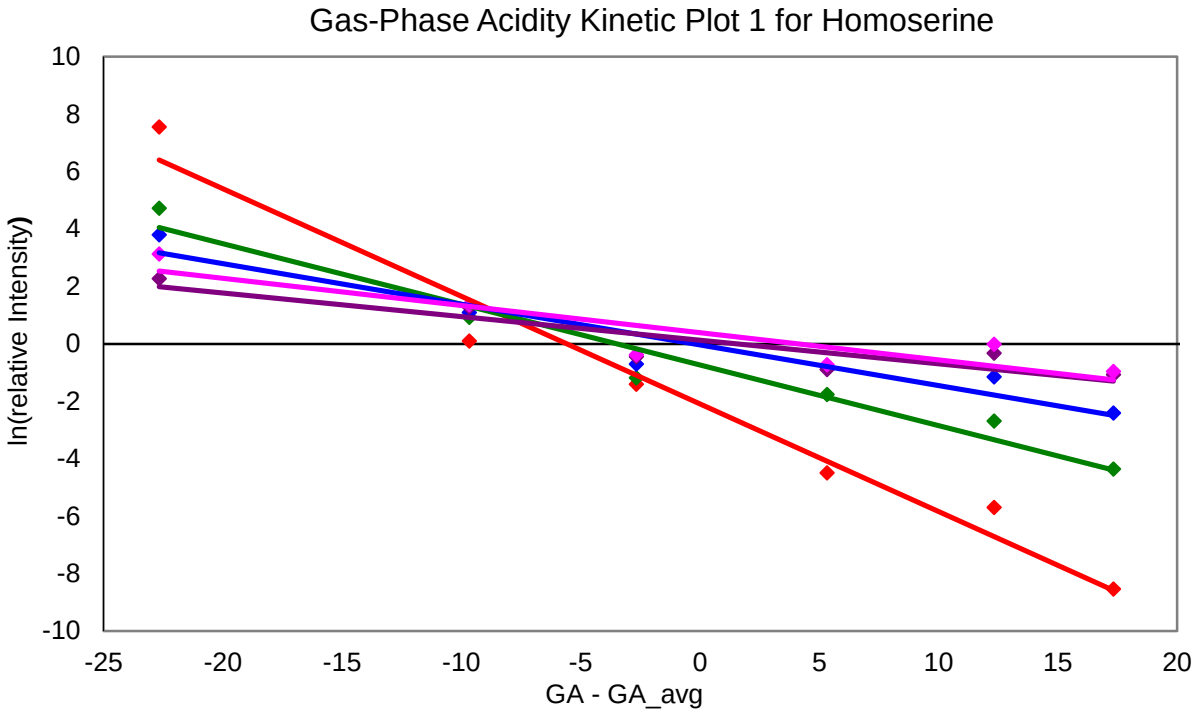


Figure 9 – Kinetic Plot 1 for Homoserine

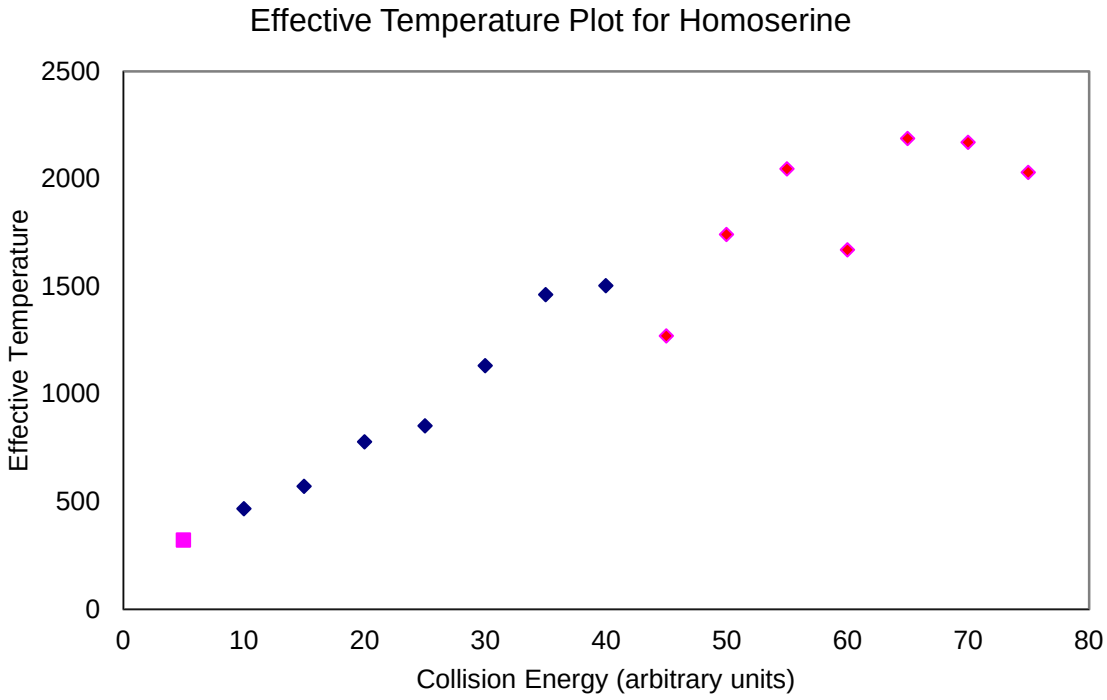


Figure 10 – Effective Temperature Plot for Homoserine

From Kinetic Plot 1, an estimation of the gas-phase acidity can be calculated. The average gas-phase acidity of all of the reference acids was calculated to be approximately 1402.7 kJ/mol. From observation, the gas-phase acidity can be estimated at 1395 kJ/mol.

As the lines do not intersect at exactly one point, a second kinetic plot is necessary. In kinetic plot 2, shown in Figure 11, the negative intercepts of each of the lines graphed are plotted against the slope of the same line. Doing this will provide a set of data, whose best fit line has a slope equal to the estimated gas-phase acidity of the compound of interest - GA_{avg}.

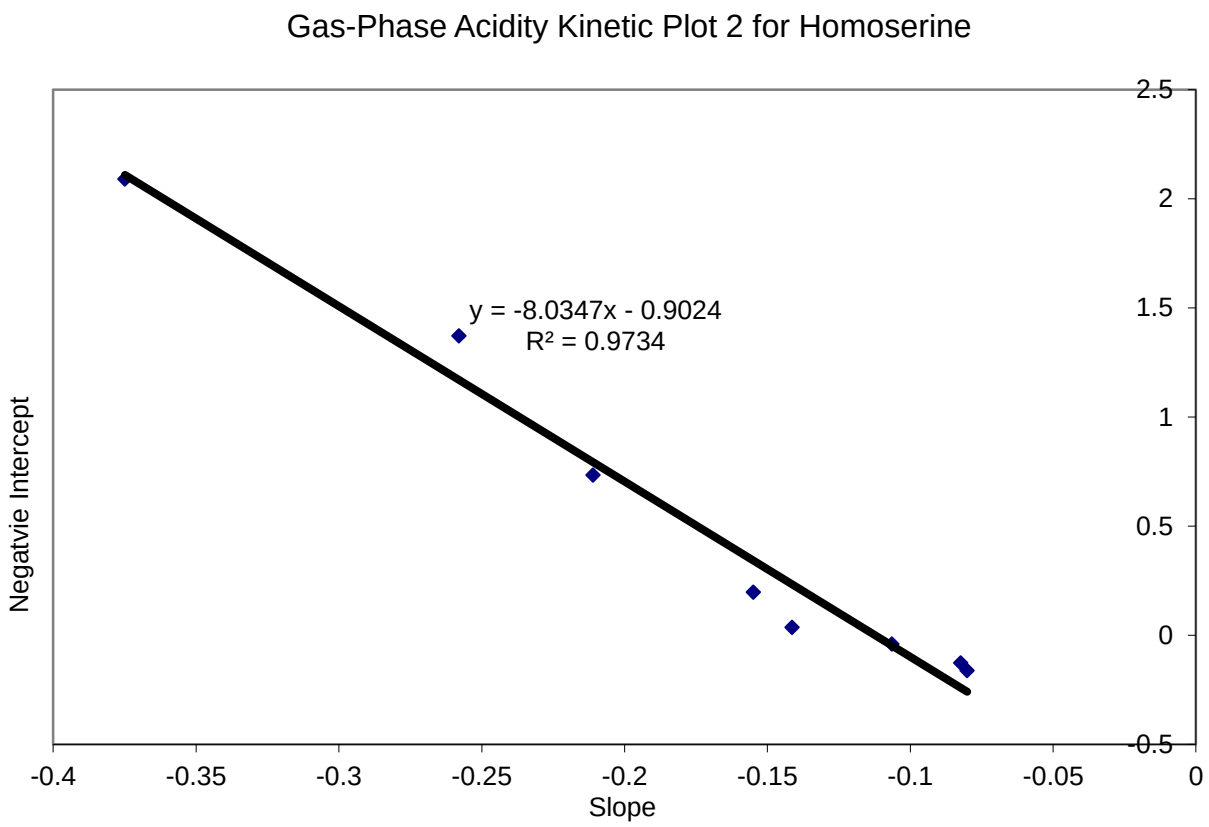


Figure 11 – Kinetic Plot 2 for Homoserine

The R^2 value of the line of best fit shows that statistical significance can be derived from this result, yielding a value of 1394.6 ± 10 kJ/mol for the gas-phase acidity of homoserine.

Thermochemical Data for Homocysteine:

In a similar manner as for homoserine, a total of five reference acids were tested against homocysteine to determine the gas-phase acidity. The five reference acids are five of the six used in determining homoserine. 3-methyl-4-nitrophenol had a gas-phase acidity that was too low compared to homocysteine's to accurately measure the relative intensities. Table 1 shows the molecular mass and gas-phase acidity for each of the five reference acids and 3-methyl-4-nitrophenol, used for homoserine, but not homocysteine. Figure 8 also shows the formulas for

each reference acid and their m/z ratio, again including 3-methyl-4-nitrophenol, though it was not used in this analysis.

Kinetic Plot 1 for homocysteine is shown in Figure 12 below. As is evident in the kinetic plot, the isothermal point is not readily obvious. Due to the very large entropic effects involved with homocysteine, compounds with similar gas-phase acidity appear too acidic to test. The natural logarithm of the relative intensities is typically above five, yielding results which are too sensitive to small changes in the ion count for homocysteine. Even with this, by extending the lines to the left a few units, the isothermal point can be guessed to be about -15 from the average. The extension of Kinetic Plot 1 is shown in Figure 13 below. Since the average gas-phase acidity of all of the reference acids was determined to be 1407.2 kJ/mol, an estimate of the gas-phase acidity for homocysteine can be given a value of approximately 1393 kJ/mol.

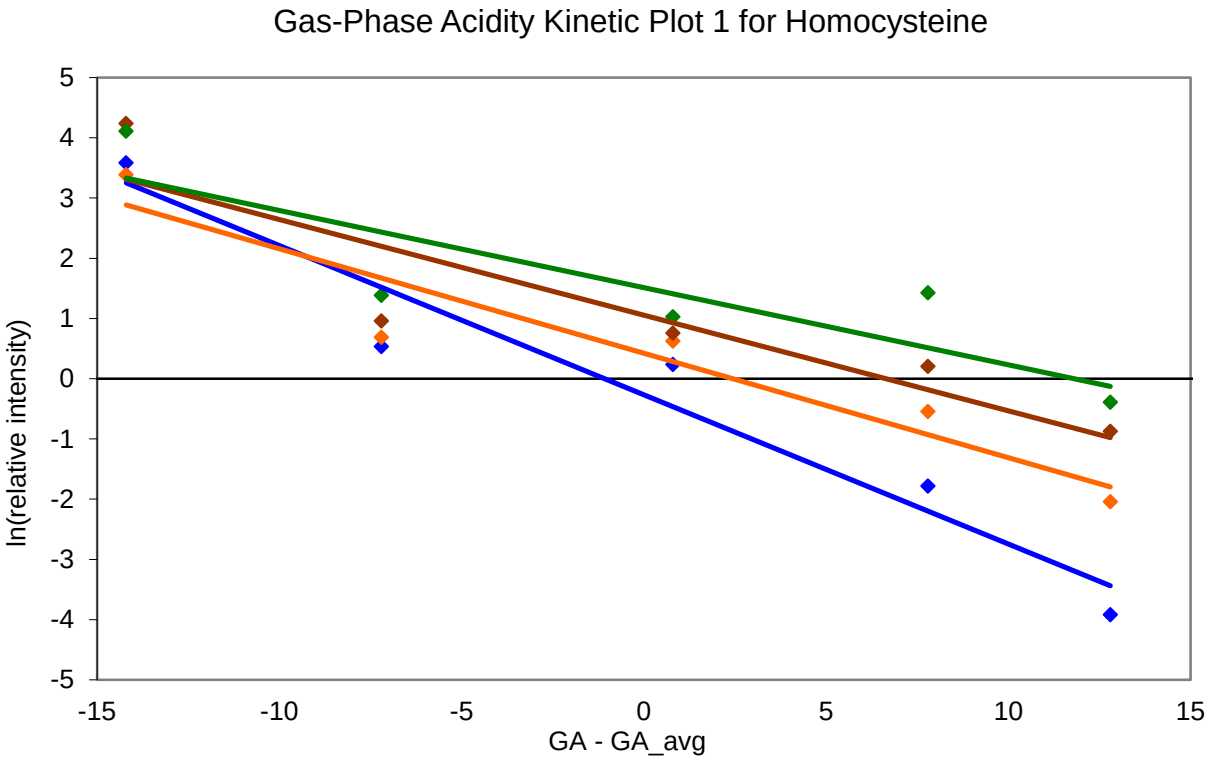


Figure 12 – Kinetic Plot 1 for Homocysteine

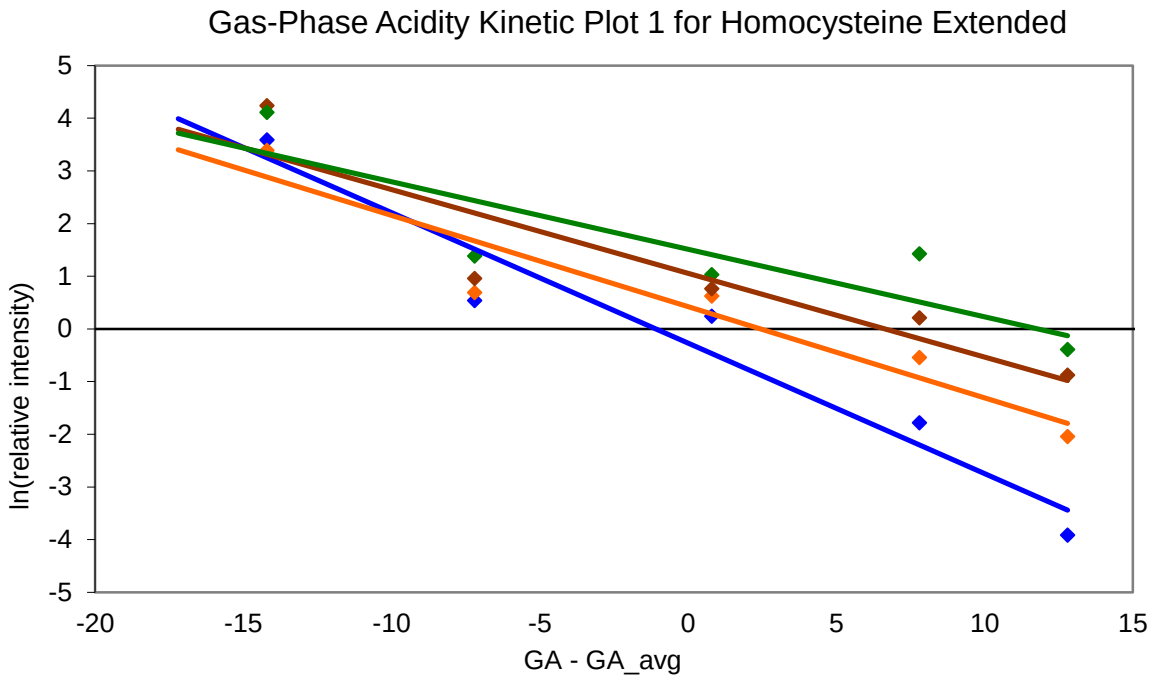


Figure 13 – Extended Kinetic Plot 1 for Homocysteine

The effective temperature plot for homocysteine is shown below in Figure 14. Unlike homoserine, the effective temperatures for homocysteine are all within reasonable ranges for effective temperatures. The blue diamonds are the energies determined to have the best values for use in the kinetic method, as they are the point of largest slope over all of the energies tested. Therefore, a total of four energies were selected from the range of 10 to 40 to be used in kinetic plot 1 above.

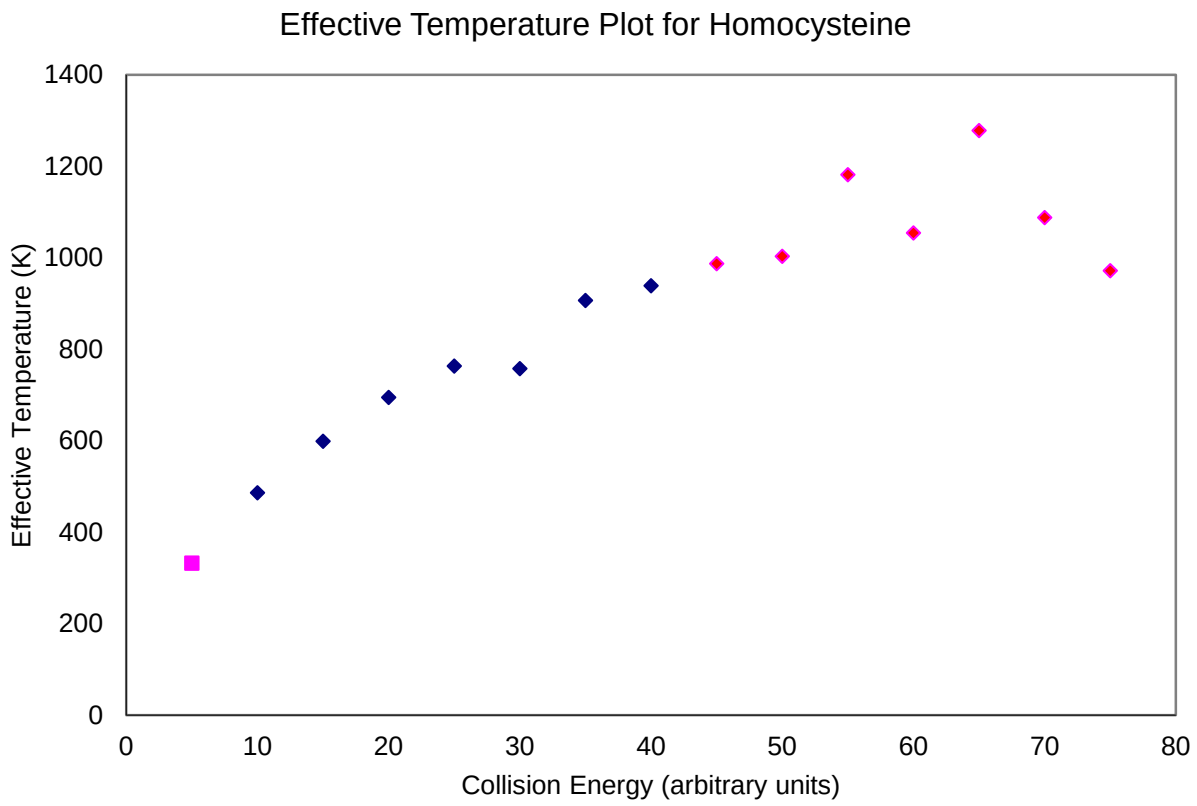


Figure 14 – Effective Temperature Plot for Homocysteine

Kinetic Plot 2 is shown below in Figure 15 for homocysteine. The R^2 value shows that there is statistical significance in the values obtained here, yielding a final value for the gas-phase acidity of homocysteine to be 1392.8 ± 10 kJ/mol.

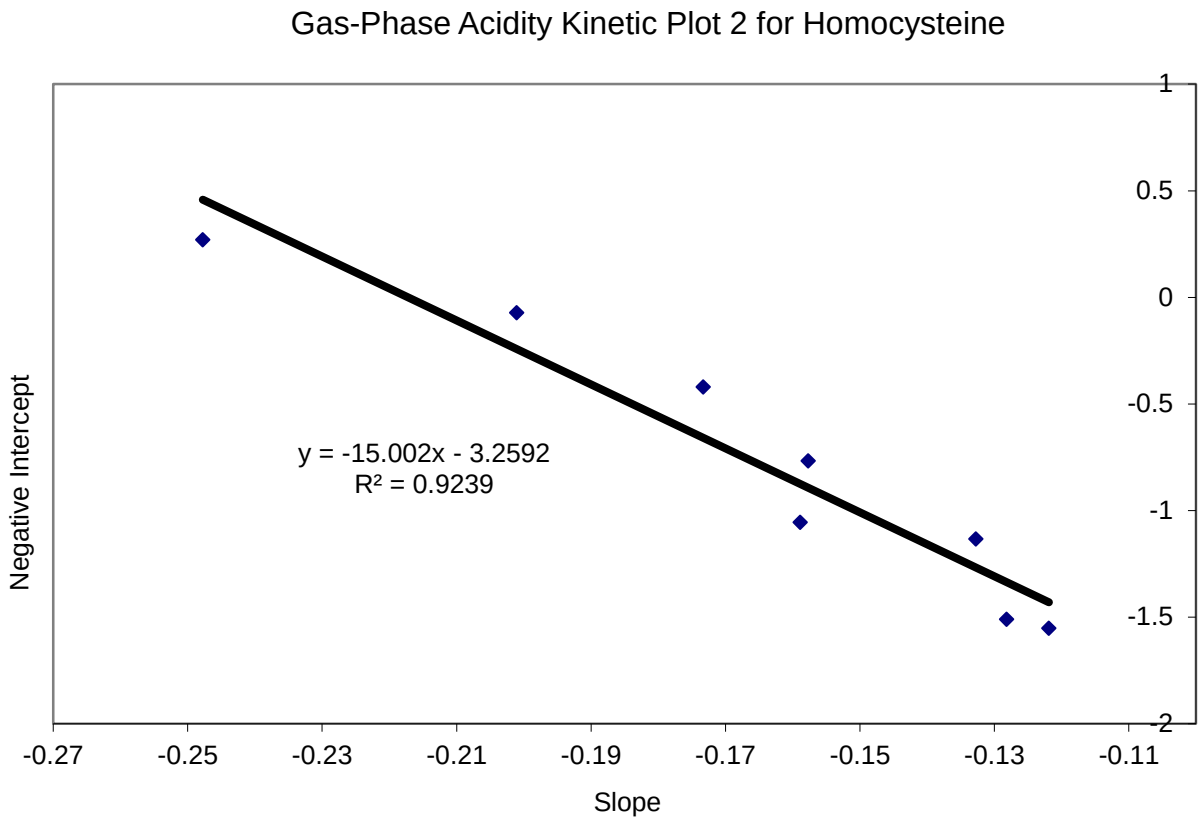


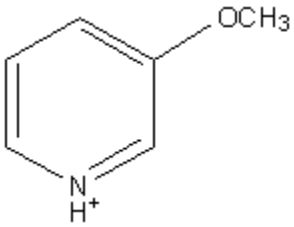
Figure 15 – Kinetic Plot 2 for Homocysteine

Thermodynamic Data for Homohomoserine:

Homohomoserine is also in the process of being investigated at the moment. Currently, the proton affinity of homohomoserine is under test with enough data to give a preliminary value for the proton affinity of homohomoserine. As this is only a preliminary value, many reference bases are still under consideration. As such, all current possible reference bases being investigated are presented here. Table 2 shows the name of each reference base, its molecular mass, and its proton affinity. Following Table 2 is Figure 16 showing the molecular formula for each of the reference bases.

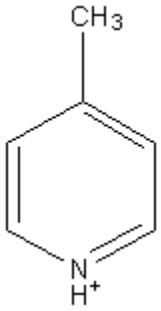
Table 2 – List of Reference Bases for Homohomoserine

Reference Base	*Molecular Mass (g/mol)*	*Proton Affinity (kJ/mol)*
3-Methoxypyridine	109	942.7
4-Picoline	93	947.2
Diethylamine	73	952.4
4-Tertbutylpyridine	135	957.7
2,3-Lutidine	107	958.9
2,4-Lutidine	107	962.9
1-Methylpyrrolidine	85	965.6
Diisopropylamine	101	971.9



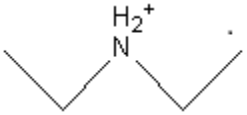
3-Methoxypyridine

m/z = 110



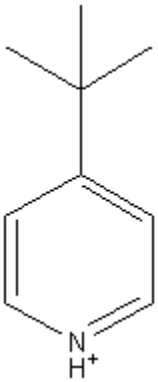
4-Picoline

m/z = 94



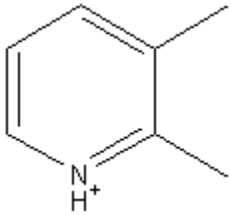
Diethylamine

m/z = 74



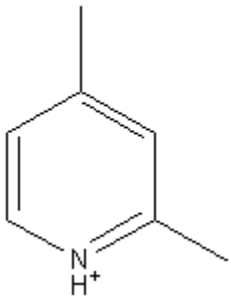
4-Tertbutylpyridine

m/z = 136



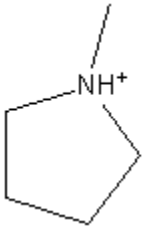
2,3-Lutidine

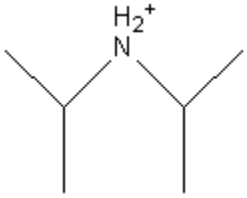
m/z = 108



2,4-Lutidine

m/z = 108





1-Methylpyrrolidine

m/z = 86

Diisopropylamine

m/z = 102

Figure 16 – Structures of Protonated Reference Bases for Homohomoserine

The mass of homohomoserine is 133, and though this is very close to the mass of 4-Tertbutylpyridine, the two mass difference is enough to establish a baseline separation between the two. As before with homocysteine, since 2,3-lutidine and 2,4-lutidine have the same mass, ample time was allowed for flushing between runs, and the runs for each reference base were performed as far apart as possible. Also, some of the proton affinities of pairs of reference base compounds are very close to one another. This is merely a technique to determine whether or not a certain reference base has any anomalies which would greatly affect the kinetic method.

Kinetic plot 1 for homohomoserine is shown in Figure 17 below. The isothermal point can be given an approximate value at the moment of 13 more than the average proton affinity, yielding an estimated proton affinity of 969 kJ/mol based off of the average proton affinity of 957.4 kJ/mol.

The effective temperature plot is shown below in Figure 18. The collision energies decided to be valid in this analysis are currently energies ranging from 5 to 30 arbitrary units. Since this is a preliminary assessment, these energies may or may not be sufficient to determine the proton affinity. For Kinetic plot 1 above, four of these energies were chosen for the graphical representation.

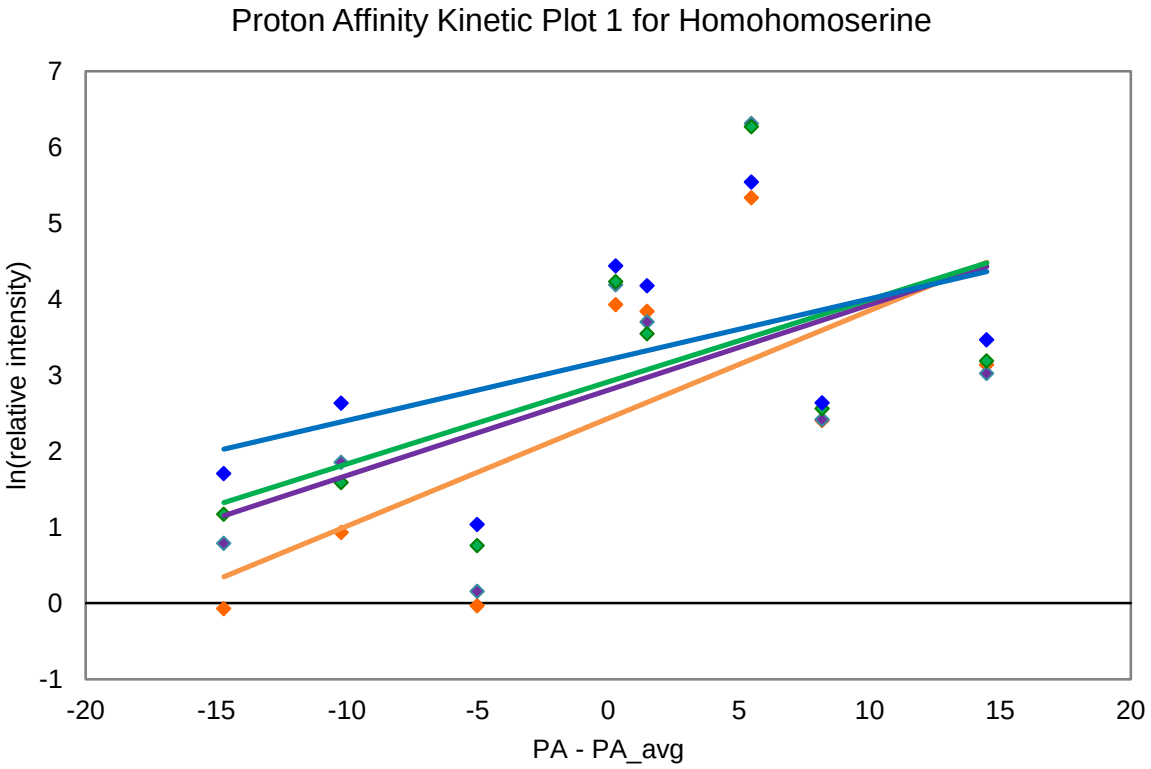


Figure 17 - Kinetic Plot 1 for Homohomoserine

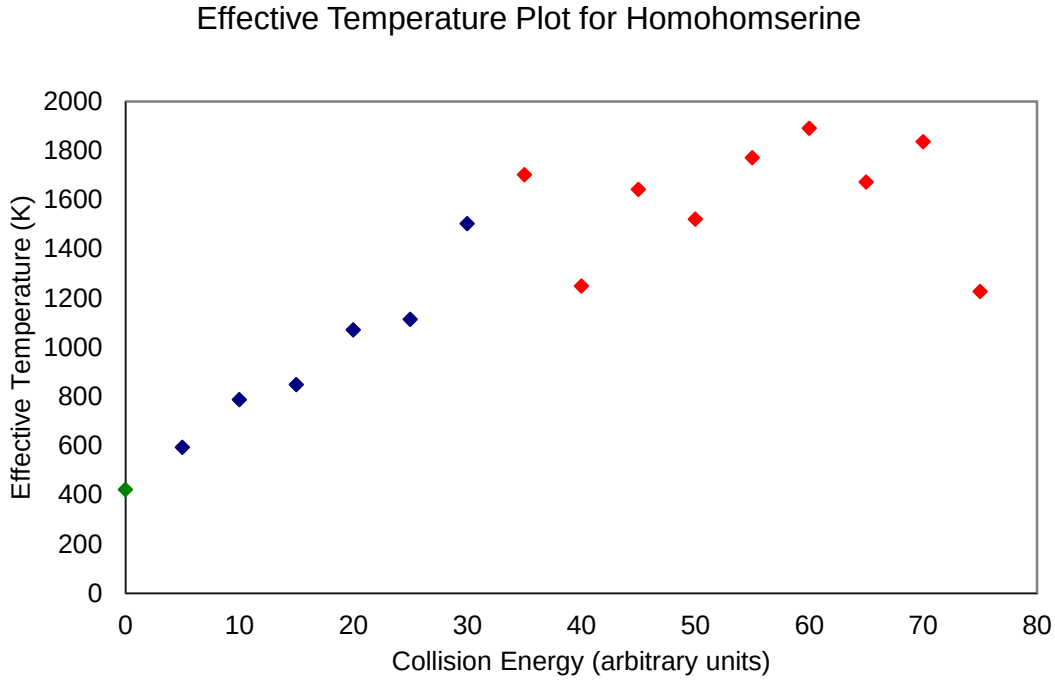


Figure 18 – Effective Temperature Plot for Homohomoserine

Kinetic Plot 2 for homohomoserine is shown below in Figure 19. The R^2 value of the line of best fit shows that there is significant statistical evidence for the value obtained here. 0The final preliminary value obtained for homohomoserine is therefore 968.7 ± 10 kJ/mol.

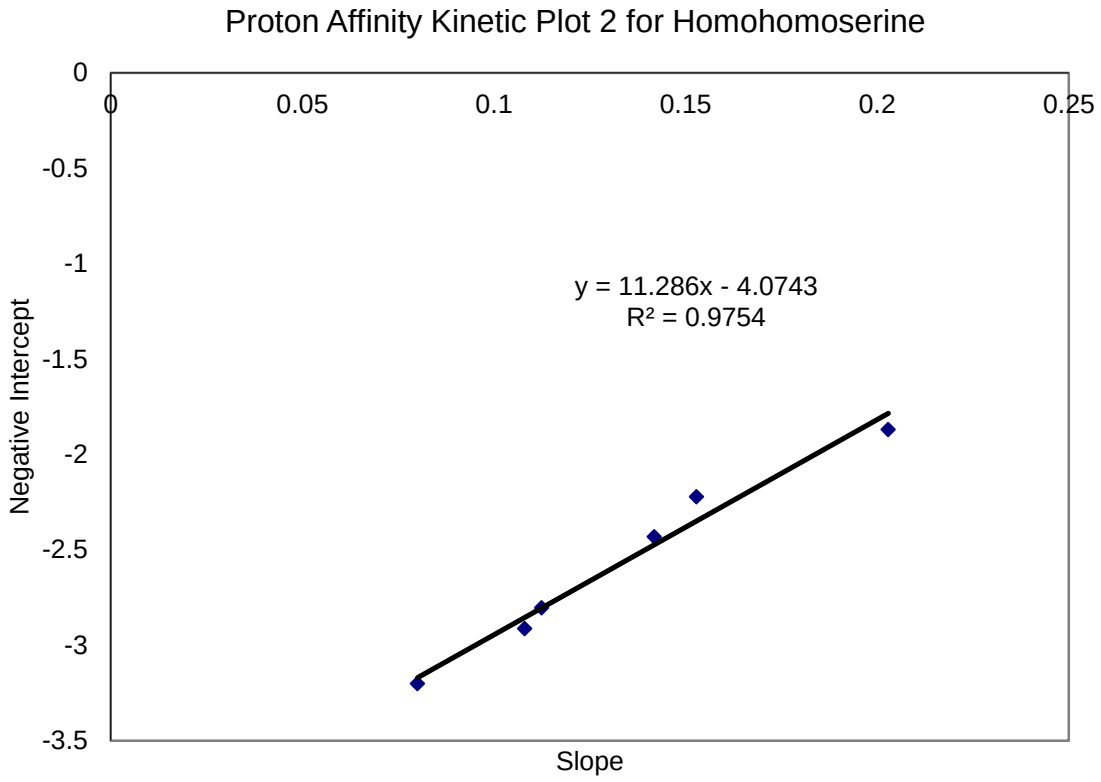


Figure 19 – Kinetic Plot 2 for Homohomoserine

Conclusions and Discussion:

Conclusions and Discussions for Serine and Its Analogues:

Table 3 below shows the compilation of thermochemical data for serine, homoserine, and homohomoserine, both experimental and theoretical values.

Table 3 – Comparison of Experimental and Theoretical Values (kJ/mol) for Serine and its Analogues

	Experimental PA	*Theoretical PA*	*Experimental GA*	*Theoretical GA*
Serine	912.5[14]	914.6[15]	1395.0 ± 9.0[18]	1392.0 ± 13.0[16]
Homoserine	942.5 ± 18.0[17]	942.1 ± 10.0[17]	**1394.6 ± 10.0**	1393.6 ± 10.0[17]
Homohomoserine	**968.7 ± 10.0**	975.0 ± 10.0[17]	-	1383.4 ± 10.0[17]

Comparing the theoretical values obtained through calculations performed earlier in this lab, the two values for serine analogues studied here are in excellent agreement with theory. The gas-phase acidity for homoserine is only about 1.0 kJ/mol off of the theoretical value, though including both of their errors makes the value obtained very reasonable. For the proton affinity of homohomoserine, the value obtained so far is still within experimental error, as it is only 6.3 kJ/mol off of the theoretical value. Again, once the error is included for both values, the number obtained is very reasonable, and should be taken as a good estimate for the true value.

In investigating the effects of the side chain on the proton affinity of serine and its analogs, it is interesting to note that the proton affinity goes up a considerable amount with each addition of a ($-CH_2$) group. The relative increase between each addition changes slightly, lowering from a difference of 40 kJ/mol to a difference of about 30 kJ/mol. The reason for this

difference is most likely that the further the electron withdrawing group (-OH) is from the protonation site, the more that the molecule "wants" the extra proton. In serine, there are only two ($-CH_2$) groups separating the (-OH) group from ($-NH_3^+$) group. As the oxygen atom is a powerful electron withdrawing group, this small distance will siphon away some of the electron density, thus destabilizing the ($-NH_3^+$) group with its already full positive charge. However, once more ($-CH_2$) groups are between the (-OH) group and the ($-NH_3^+$) group, the amount of electron density shifted away from positively charged amine group goes down, thus further stabilizing this protonated molecule. A potential reason for the decreases in the difference between the analogues is that in homohomoserine, the (-OH) withdraws less electron density than the (-OH) group in homoserine, but because the (-OH) group is already a fair distance away from the ($-NH_3^+$) group, it cannot withdraw the same magnitude of density as it originally could going from serine to homoserine.

One effect that is also noticeable is the fact that an intramolecular hydrogen bond is formed in the protonated amino acids. When serine and its analogues are protonated, the amine group and the hydroxyl group form an intramolecular hydrogen bond.[17] This intramolecular hydrogen bond creates a ring system which adds to the discussion of stability of the ion. In serine, a five-membered ring is formed, which is stable, though has some added strain. In homoserine, the ring size is now six, which is very stable, adding to the proton affinity. In homohomoserine, the ring size is seven. This addition is still stable, especially more so than normal serine. These additions help stabilize the higher forms even more than serine. This analysis is supported by comparing it to the work done on lysine analogues.[13] Here, it was shown that by reducing the length of the side chain, the proton affinity went down considerably. Lysine also forms an intramolecular hydrogen bond when protonated, and has a ring size of eight. Its

three other analogues have ring sizes of seven, six, and five. The proton affinities of these forms also decrease uniformly, giving more evidence that the ring size of the intramolecular hydrogen bond affects the overall stability of the protonated molecule.

An important distinction should be noticed in the gas-phase acidity of serine and its analogues. Though experimental values are not shown, comparing the experimental acidity obtained in this research with the theoretical obtained from sources shows a very different situation from that of proton affinity. The values do not vary significantly with the addition of the extra ($-CH_2$) groups. When accounting for the error that is possible in all three values, the ordering of the three is no longer certain. As the only major difference between the three molecules is the distance that the electron withdrawing group (-OH) is from the site of deprotonation, this distance apparently does not affect the stabilization of the ion. This makes sense in this case, due to the large stability of the ($-CO_2^-$) group. This group is very stable due to the small πsystem available for the electrons to travel. As this π system already stabilizes the ion, any additional electron withdrawing nature gained from having a close (-OH) group is overshadowed.

Conclusions and Discussions for Cysteine and its Analogues:

Shown in Table 4 below are the calculated and experimental values for the gas-phase acidities of cysteine and its analogues. The theoretical value is shown both for its carboxylate anion and thiolate anion forms

Table 4 – Comparisons of Experimental and Theoretical Values (kJ/mol) for Cysteine and its Analogues

	Experimental GA	*Theoretical GA carboxylate*	*Theoretical GA thiolate*
Cysteine	1392.9 ± 13.8[11]	1394.0 ± 10.0[18]	1386.0 ± 10.0[17]
Homocysteine	**1392.8 ± 10**	1398.8 ± 10.0[17]	1397.2 ± 10.0[17]
Homohomocysteine	-	1412.0 ± 10.0[17]	1408.5 ± 10.0[17]

Looking at how the value obtained in this experiment for the gas-phase acidity of homocysteine compared to the two theoretical values, both of them are possible. Given the large error ranges associated with all three values, no one form can be stated as being the predominate form with any certainty. However, as was the analysis done by Janiga, it appears as if the thiolate form dominates the carboxylate form for cysteine, homocysteine, and homohomocysteine.[17]

Regardless of the form of the ion, the data suggests that an increase of some sort is occurring between the three forms of cysteine for the gas-phase acidity. This increase must be attributed to the increased separation of the (-SH) group from the ($-CO_2H$) group, as the trend is observed for both the carboxylate and the thiolate ion forms. One possible explanation for this effect would be that the two groups destabilize the ion when they are too close to one another. Since the acidity of the molecule goes up as the distance between the two groups increase, this is a feasible explanation for the increase.

Future Work:

Future work on this subject is expected to follow a simple path. First, the proton affinity value for homohomoserine needs to be finished, then the proton affinity value for homohomocysteine will be attempted. One possible issue with the testing of homohomocysteine is the form in which it currently is. The only source of pure homohomocysteine available at the

moment is the homohomocysteine molecule with a disulfide bond already formed. A simple acidic solution should break this bond, though more complications may arise with the gas-phase acidity analysis for homohomocysteine. As the solution must be basic to run the kinetic method for gas-phase acidity and combining that with the facts that homohomocysteine tends to form a disulfide bond when in a basic solution and only a small sample is available, considerable difficulties are expected. Similarly, the homohomoserine solution that would be needed to be prepared would also need to be basic, which should not be too difficult, as no unfortunate reactions seem to occur in basic solutions for serine homologues. After these values are tabulated, a more rigorous analysis of the effect of the side chain on gas-phase thermochemistry can be conducted, as well as a better understanding of the deprotonation site for cysteine and its analogues in the gas-phase.

References

[1] Dempster, A. J. (1918). "A New Method of Positive Ray Analysis". *Phys. Rev.* **11**: 316–325.
[2] Snyder, A. P. (1996). "Electrospray: A Popular Ionization Technique for Mass Spectrometry". *ACS Symposium Series* **619**: 1-20.
[3] Bowers, M. T, et al. (1996). "Mass Spectrometry: Recent Advances and Future Directions". *J. Phys. Chem* **31**: 12897-12910.
[4] Bruins, A. P. (1991). "Mass spectrometry with ion sources operating at atmospheric pressure". *Mass Spectrometry Reviews* **10**: 53–77.
[5] Buchana, M. V. and Comisarow, M.B (1987). "Principles and Features of Fourier Transform Mass Spectrometry". *ACS Symposium Series* **359**: 1-20.
[6] Carter, J. A. et al. (1975). "Trace Elements in Fuels by Isotope Dilution Mass Spectrometry". *Advances in- Chemistry* **141**: 74-83.
[7] Burlingame, K. F. et al. (1996). "From Protein Primary Sequence to the Gamut of Covalent Modifications Using Mass Spectrometry". *Mass Spectrometry* **24:** 472-511.
[8] Ervin, K. M. (2001). "Experimental Techniques in Gas-Phase Ion Thermochemistry". *Chem. Rev.* **2**: 391-444.
[9] Tsang, Yuet et. al. (2004). "Experimental validation of theoretical potassium and sodium cation affinities of amides by mass spectrometric kinetic method measurements." *Rapid Commun. Mass Spectrom* **18**: 345-355.
[10] Polfer, N. C. et al. (2005) "Infrared Fingerprint Spectroscopy and Theoretical Studies of Potassium Ion Tagged Amino Acids and Peptides in the Gas Phase". *J. Am. Chem. Soc.* **23**: 8571-8579
[11] Tian, Z. et al. (2007). "Are Carboxyl Groups the Most Acidic Sites in Amino Acids? Gas-Phase Acidity, H/D Exchange Experiments, and Computations on Cysteine and Its Conjugate Base". *J. Am. Chem. Soc.* **17**: 5403-5407.
[12] Wesdemiotis, C. *J.* (2004) *Mass Spectrom.* **39**: 998-1003
[13] Schroeder, O. E., Andriole, E.J., Carver, K.L., Colyer, K.E., Poutsma, J.C. (2004)"Proton Affinity of Lysine Homologues from the Extended Kinetic Method". *J. Phys. Chem*, **108**: 326-332
[14] Gronert. S., et. al. (2009) "A Reevalutaion of Computed Proton Affinities for the Common α-Amino Acids". *J. Am. Soc. Mass. Spectrom.* **20**: 2116-2123.
[15] Hunter, E. P. L. and Lias, S. G. (1998) "Evaluated Gas Phase Basicities and Proton Affinities of Molecules: An Update". *J. Phys. Chem. Ref. Data* **27**: 413-656.
[16] Bleiholder, C. et. al. (2006) "Revising the Proton Affinity Scale of the Naturally Occurring α-Amino Acids". *J. Am. Soc. Mass. Spectrom.* **17**: 1275-1281.
[17] Janiga, A. R. (2010) "Computational and Experimental Investigations of Gas-Phase Thermodynamics of Non-Protein Amino Acids". MS Thesis at The College of William and Mary.
[18] Jones, M.C. et al (2007) "Gas-Phase Acidities of the 20 Protein Amino Acids". *Int. J. Mass Spectrom.* **267**: 54-62